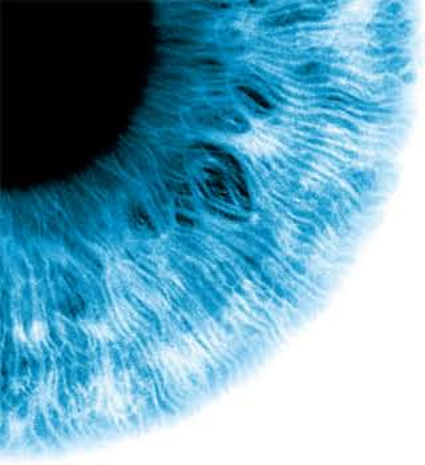


# Common mental health disorders: Evidence Update March 2013

**A summary of selected new evidence relevant to NICE clinical guideline 123 'Common mental health disorders: identification and pathways to care' (2011)**

Evidence Update 31

Evidence Updates provide a summary of selected new evidence published since the literature search was last conducted for the accredited guidance they relate to. They reduce the need for individuals, managers and commissioners to search for new evidence. Evidence Updates highlight key points from the new evidence and provide a commentary describing its strengths and weaknesses. They also indicate whether the new evidence may have a potential impact on current guidance. For contextual information, this Evidence Update should be read in conjunction with Common mental health disorders (NICE clinical guideline 123).

**Evidence Updates do not replace current accredited guidance and do not provide formal practice recommendations.**

NHS Evidence is a service provided by NICE to improve use of, and access to, evidence-based information about health and social care.

**National Institute for Health and Clinical Excellence**

Level 1A
City Tower
Piccadilly Plaza
Manchester M1 4BT
www.nice.org.uk

# Contents

# Introduction

This Evidence Update identifies new evidence that is relevant to, and may have a potential impact on, the following reference guidance:

[1] **Common mental health disorders. NICE clinical guideline 123 (2011).**

A search was conducted for new evidence from 1 September 2010 to 15 October 2012. A total of 5305 pieces of evidence were initially identified. Following removal of duplicates and a series of automated and manual sifts, 9 items were selected for the Evidence Update (see Appendix A for details of the evidence search and selection process). An Evidence Update Advisory Group, comprising topic experts, reviewed the prioritised evidence and provided a commentary.

Although the process of updating NICE guidance is distinct from the process of an Evidence Update, the relevant NICE guidance development centres have been made aware of the new evidence, which will be considered when guidance is reviewed.

## Other relevant NICE guidance

The focus of the Evidence Update is on the guidance stated above. However, overlap with other accredited guidance has been outlined as part of the Evidence Update process. Where relevant, this Evidence Update therefore makes reference to the following guidance:

- [2] Antenatal and postnatal mental health. NICE clinical guideline 45 (2007).

## Feedback

If you have any comments you would like to make on this Evidence Update, please email contactus@evidence.nhs.uk

[1] NICE-accredited guidance is denoted by the Accreditation Mark
[2] Guidance published prior to NICE accreditation

# Key points

The following table summarises what the Evidence Update Advisory Group (EUAG) decided were the key points for this Evidence Update. It also indicates the EUAG's opinion on whether the new evidence may have a potential impact on the current guidance listed in the introduction. For further details of the evidence behind these key points, please see the full commentaries.

The section headings used in the table below are taken from the guidance.

**Evidence Updates do not replace current accredited guidance and do not provide formal practice recommendations.**

| | Potential impact on guidance | |
|---|---|---|
| **Key point** | **Yes** | **No** |
| **Improving access to services** | | |
| • Limited evidence suggests that a community-based intervention of rapid access to care may result in improved mental health in black African or black Caribbean populations. | | ✓ |
| • A Malaysian translation of the 7-item Generalized Anxiety Disorder (GAD-7) scale seems to be reliable and correlates with results of other diagnostic instruments. | | ✓ |
| **Step 1: identification and assessment** | | |
| • Asking 2 general case-finding questions about depression may be useful for identifying or ruling out postnatal depression. | | ✓ |
| • Generally, instruments for identifying anxiety and depression in people with learning disabilities seem to be insufficiently studied, although the Glasgow Anxiety Scale for people with an Intellectual Disability (GAS-ID), the Anxiety Depression and Mood Scale (ADAMS) and the Glasgow Depression Scale for people with a Learning Disability (GAS-LD) may be useful. | | ✓ |
| • GPs may correctly rule out distress and depression in about 80% of people who do not have distress or depression; however GPs may only diagnose distress correctly in about half of people with distress and may only diagnose depression correctly in about a third of people who have depression. | | ✓ |
| • The 15-item Geriatric Depression Scale (GDS) may be useful in primary care for diagnosing depression in older people. | | ✓ |
| • The Patient Health Questionnaire (PHQ)-9, using a diagnostic cut-off score of 10, may have increased sensitivity in primary care compared with hospital settings. | | ✓ |
| **Steps 2 and 3: treatment and referral for treatment** | | |
| • Evidence for telephone counselling in people with acquired physical disabilities seems to be limited, but may indicate beneficial effects on coping strategies, community integration and depression. | | ✓ |

# 1 Commentary on new evidence

These commentaries analyse the key references identified specifically for the Evidence Update. The commentaries focus on the 'key references' (those identified through the search process and prioritised by the EUAG for inclusion in the Evidence Update), which are identified in bold text. Supporting references provide context or additional information to the commentary. Section headings are taken from the guidance.

## 1.1 Improving access to services

**Community-based rapid care intervention**

NICE CG123 recommends that primary and secondary care clinicians, managers and commissioners should collaborate to develop local care pathways that promote access to services for people with common mental health disorders by: supporting the integrated delivery of services across primary and secondary care; having clear and explicit criteria for entry to the service; focusing on entry and not exclusion criteria; having multiple means (including self-referral) to access the service; and providing multiple points of access that facilitate links with the wider healthcare system and community in which the service is located.

**Afuwape et al. (2010)** reported on a randomised UK study of a community-based intervention in 40 people of black African or black Caribbean origin meeting World Health Organization (WHO) criteria for anxiety or depression. Recruitment was by both self-referral and referral from statutory and non-statutory organisations. Additionally, study staff attempted to identify people in key organisations such as local businesses and the church that could advocate on behalf of the service and encourage people in their communities to contact the service.

People underwent a simple screen for eligibility before assessment by a member of staff and a researcher using the WHO mental health checklist, and then baseline assessments were performed before randomisation. People assigned to the rapid care group were contacted and began treatment; those in the control group were offered an appointment in 3 months.

The rapid care intervention consisted of a needs-led care programme including practical advice and assistance, advocacy for social needs, health education and mentoring, and brief therapies based on principles of cognitive behaviour therapy and solution-focused therapy. The primary outcome measure was general mental health assessed by the General Health Questionnaire (GHQ)-28 at 3 months.

Differences in groups at baseline were seen: people in the rapid care group were older, with less previous psychiatric history and higher GHQ-28 scores, although the authors recognised that baseline differences were likely to be due to the small sample size.

After adjustment for baseline GHQ-28 score, psychiatric history and age, the mean difference was 7.76 (95% confidence interval [CI] 0.86 to 14.65) with better mental health seen in the rapid care group. Mean costs were £915 (standard deviation [SD] £997) in the rapid care group and £912 (SD £906) in the control group. After adjustment for baseline costs, rapid care cost £274 more than control (bootstrapped 90% CI −£176 to £740) but this difference was not significant.

The randomised sample size of 20 people per group was substantially lower than the 42 people per group stated in the sample size calculation, and only 16 people in each group completed the study. The authors also noted that although engagement with local organisations was slow, 37% of all contacts were either self-referrals or referred by a friend. In addition to the limitations identified by the authors, the use of a waiting list control rather than usual care had potential to bias results.

Further UK-based studies like this are needed to establish an evidence base for different formats and locations for therapy in common mental health disorders. This limited evidence provides some support for local care pathways that facilitate links with the wider healthcare system and community in which the service is located, as recommended in NICE CG123.

**Key reference**

Afuwape SA, Craig TKJ, Harris T et al. (2010) The Cares of Life Project (CoLP): An exploratory randomised controlled trial of a community-based intervention for black people with common mental disorder. Journal of Affective Disorders 127: 370–4

### Validation of a Malaysian translation of GAD-7

NICE CG123 recommends considering modifications to the method and mode of delivery of assessment and treatment interventions and outcome monitoring (based on an assessment of local needs), which may typically include using: technology (for example, text messages, email, telephone and computers) for people who may find it difficult to, or choose not to, attend a specific service; and bilingual therapists or independent translators.

**Sidik et al. (2012)** reported on a study (n=895) to validate the Malay version of the 7-item Generalized Anxiety Disorder (GAD)-7 scale. This study was part of a larger study of the prevalence of depression and anxiety in women and validity of Malay versions of several questionnaires on depression and anxiety. The process for translation followed guidelines for cross-cultural adaptation of self-report measures and consisted of 2 independent forward translations of GAD-7 into Malay, then reaching consensus between translators on the forward translation, back translation and review of back translation by an expert committee.

The results obtained using GAD-7 were compared with scores for GHQ-12, which does not specifically detect depression or anxiety but gives an overall estimate of the prevalence of psychiatric morbidity. The Composite International Diagnostic Interview (CIDI) anxiety module was used as the reference standard.

Malaysian women aged 18 years and older who did not have acute illness needing immediate medical attention or communication difficulties were recruited consecutively by research assistants. Participants then completed a self-administered questionnaire that included GAD-7, GHQ-12, and the Patient Health Questionnaire (PHQ)-9; although the results of the PHQ-9 validation study were not part of this report. Participants were divided into 2 groups: normal scores (PHQ-9<10 and GAD-7<5) and high scores (PHQ-9≥10 or GAD-7≥5).

Systematic weighted random sampling was conducted to select participants from each group for interviewing with the CIDI software, which was used to generate a diagnosis based on Diagnostic and Statistical Manual of Mental Disorders (DSM)-IV criteria. People who had known psychiatric illness or were taking psychoactive drugs were excluded from the data analysis; the authors stated that excluding these participants during data collection was not possible.

Overall, 82% of participants (n=730) had normal scores in both GAD-7 and PHQ-9 and 18% (n=165) had high scores in one or both questionnaires. The authors reported that the Malay version of GAD-7 had good internal reliability (Cronbach's alpha=0.74, a value of 0.7 or more was noted to indicate that a questionnaire is reliable). The Pearson's correlation coefficient for GAD-7 and GHQ-12 was 0.62, indicating a positive association of moderate strength. Against the CIDI anxiety module, GAD-7 had sensitivity of 76.3% (95% CI 60.8% to 87.0%), specificity of 94.4% (95% CI 88.4% to 97.4%), positive likelihood ratio of 13.74 (95% CI 6.19 to 30.50), and negative likelihood ratio of 0.25 (95% CI 0.14 to 0.45).

A limitation discussed by the authors was that time and resource constraints led to conducting the study in only 1 urban government-funded clinic. The participants were women of low to middle income, so may not represent all Malaysian women. Additionally, these results may not be applicable to Malaysian men.

Although the use of the Malay translation of GAD-7 in the UK is likely to be low, this study shows an example of using a recommended technique to translate a diagnostic questionnaire into another language to retain the validity of the tool. This evidence is not likely to affect recommendations in NICE CG123 to consider using bilingual therapists or independent translators, but clinicians should be aware that simple translation of tools may not be as effective as using a validated translation.

**Key reference**

Sidik SM, Arroll B, Goodyear-Smith F (2012) Validation of the GAD-7 (Malay version) among women attending a primary care clinic in Malaysia. Journal of Primary Health Care 4: 5–11

## 1.2 Stepped care

No new key evidence was found for this section.

## 1.3 Step 1: identification and assessment

**Case-finding in postnatal depression**

NICE CG123 recommends asking 2 questions in people who may have depression: 'During the last month, have you often been bothered by feeling down, depressed or hopeless?' and 'During the last month, have you often been bothered by having little interest or pleasure in doing things?' If the person answers 'yes' to either of these questions further assessment for depression should be considered. A similar recommendation in 'Antenatal and postnatal mental health' (NICE CG45) states: If the woman answers 'yes' to either of the initial questions, also ask: 'Is this something you feel you need or want help with?' In the postnatal population, a 'yes' response to this third question would lead to further assessment for depression.

**Mann and Gilbody (2011)** did a systematic review of randomised controlled trials and prospective cohort studies that used 2 questions for case-finding as a method of recruitment in women within 1 year of a live birth and were assessed for non-psychotic depressive disorder in any country and in any setting. The questions were those recommended in NICE CG123; studies using PHQ-2 were excluded. PHQ-2 consists of the same two questions but is framed over the previous 2 weeks and is scored on a scale from 0 (not at all) to 3 (nearly every day).

Studies were eligible for inclusion if postnatal depression was diagnosed using DSM-IV, International Classification of Disease (ICD)-10 or Research Diagnostic Criteria (RDC). However, only 1 diagnostic test accuracy study comparing the 2 case-finding questions with a gold-standard reference test in postnatal women was identified. Six studies were excluded because they did not use a gold-standard reference test.

The 1 included study (n=506) reported a 100% sensitivity (95% CI 79% to 100%) at up to 1 month after birth, suggesting that negative responses to the case-finding questions may rule out depression. The specificity at up to 1 month after birth was 62% (95% CI 57% to 66%); the probability that positive responses indicated postnatal depression was 11% (95% CI 7% to 17%), suggesting that the case finding questions may not be enough to diagnose postnatal depression.

Quality assessment indicated that this study may have had incorporation bias and review bias. The authors noted that further research to assess the diagnostic test accuracy of the 2 case-finding questions is needed.

Although this review shows that the evidence-base for diagnostic accuracy studies is limited, the available evidence supports the recommendation in NICE CG123 to use two specific case-finding questions in people with suspected depression, and these questions may also be useful in suspected postnatal depression. Additionally, the need to seek further assessment for depression if a person has positive responses is supported by this evidence.

Additional information about the study by Mann and Gilbody (2011) is also available from an independent critical appraisal report produced for the Centre for Reviews and Dissemination's Database of Abstracts of Reviews of Effects.

**Key reference**

Mann R, Gilbody S (2011) Validity of two case finding questions to detect postnatal depression: a review of diagnostic test accuracy. Journal of Affective Disorders 133: 388–97

**Supporting reference**

Centre for Reviews and Dissemination (2012) Validity of two case finding questions to detect postnatal depression: a review of diagnostic test accuracy. Database of Abstracts of Reviews of Effects

### Identifying common mental health disorders in people with learning disabilities

NICE CG123 recommends that for people with significant language or communication difficulties, for example people with sensory impairments or a learning disability, clinicians should consider using the Distress Thermometer and/or asking a family member or carer about the person's symptoms to identify a possible common mental health disorder. If a significant level of distress is identified, offer further assessment or seek the advice of a specialist.

**Hermans et al. (2011)** conducted a systematic review of studies of characteristic and psychometric properties of diagnostic tools for anxiety in people with learning disabilities. Studies were included if: they consisted of at least 30 participants; psychometric properties were studied separately for the anxiety subscale; and 50% or more of participants had learning disabilities and were aged 16 years or older.

Quality of studies was assessed by a modified version of Quality Assessment of Diagnostic Accuracy Studies (QUADAS) criteria. These assessments were then graded by the authors as 'inferior', 'insufficient', 'sufficient', 'good' or 'excellent'. Of 17 studies eligible for review, the methodological quality was assessed as: 'insufficient' in 4; 'sufficient' in 7; and 'good' in 6. The studies covered 11 screening instruments (4 self-report and 7 informant-report) and 3 diagnostic instruments. The number of participants was not reported.

The authors concluded that the Glasgow Anxiety Scale for people with an Intellectual Disability (GAS-ID) was the most promising self-report instrument, on the basis of 1 study of 'sufficient' quality. This study showed 'excellent' internal consistency of 0.96, 'very high' test–retest correlation of 0.95, and 'high' correlation of 0.75 with the Beck Anxiety Scale. Using a diagnostic cut-off of 13–15 resulted in sensitivity of 100% (95% CI 83% to 100%) and specificity of 100% (95% CI 82% to 100%). Other self-report screening instruments did not have complete validity and reliability data, and available correlation scores were 'moderate' or 'low'.

For informant-report instruments, complete data were not available for any test. The 'most promising' informant-report instrument was the Anxiety, Depression and Mood Scale (ADAMS). Internal consistency was 'good' (mean of 3 results from 2 studies=0.83) and test-retest intraclass correlation coefficient was 'good' (mean of 3 results from 2 studies=0.87).

Generally, the internal consistency or reliability of other instruments was lower, or the study was of low methodological quality.

In a similar study, **Hermans and Evenhuis (2010)** systematically reviewed evidence for the feasibility, reliability and validity of instruments for assessing depression in people with learning disabilities. 24 studies of 15 instruments were identified; the number of participants was not clearly reported.

The authors concluded that the 'most promising' self-report instrument was the Glasgow Depression Scale for people with a Learning Disability (GDS-LD), which had 'excellent' internal consistency of 0.90; 'very high' test-retest correlation of 0.97 (95% CI 0.96 to 0.99); using a diagnostic cut-off score of 15, it had sensitivity of 90% (95% CI 77% to 95%) and specificity of 100% (95% CI 83% to 100%); it also had 'excellent' correlation with the Beck Depression Inventory (BDI)-II. Other tests showed poorer or conflicting results or were insufficiently studied. For informant-report instruments, although some showed promise, further studies are needed.

The authors noted that many studies were judged to be low quality because of incomplete or unclear data presentation or unclear reporting of sample characteristics. Additionally outcome measures such as sensitivity and specificity were often lacking.

Hermans et al. (2011) concluded that future research should study the psychometric properties of existing instruments to detect anxiety in people with learning disabilities, rather than developing new instruments, and that clinicians should take care in interpreting results of instruments that have not been fully studied. Neither this evidence nor that of Hermans and Evenhuis (2012) is likely to affect NICE CG123.

**Key references**

Hermans H, van der Pas FH, Evenhuis HM (2011) Instruments assessing anxiety in adults with intellectual disabilities: a systematic review. Research in Developmental Disabilities 32: 861–70

Hermans H, Evenhuis HM (2010) Characteristics of instruments screening for depression in adults with intellectual disabilities: systematic review. Research in Developmental Disabilities 31: 1109–20

### Identifying distress and mild depression in primary care

NICE CG123 recommends that if the identification questions indicate a possible common mental health disorder, but the practitioner is not competent to perform a mental health assessment, the person should be referred to an appropriate healthcare professional. If this professional is not the person's GP, the GP should be informed of the referral.

**Mitchell et al. (2011)** did a meta-analysis of studies of detection of distress or mild depression in general practice. Data for unassisted clinical diagnoses were included: if a study used an intervention to improve diagnosis, only data obtained before the intervention occurred was used. The meta-analysis aimed to find the prevalence of distress in primary care, and to assess the ability of GPs to detect distress and mild depression. 16 studies with sensitivity and specificity data for detecting distress (n=13,993) and 5 for mild depression were identified.

The adjusted prevalence of distress was 44.1% (95% CI 34.7% to 53.8%) and that of mild depression was 10.6% (95% CI 6.8% to 15.1%; typographical error in journal article corrected by Mitchell AJ, personal communication 2013). GPs correctly identified 48.4% (95% CI 42.6% to 54.2% of people with mild distress and 79.4% (95% CI 74.3% to 84.1%) of people without distress. GPs also correctly identified 33.8% (95% CI 27.3% to 40.7%) of people with mild depression, and 80.6% (95% CI 66.4% to 91.6%) of people without depression.

The authors noted several limitations of their study including that studies mostly used a single assessment to detect distress or depression, which may not reflect clinical practice and clinicians were often aware of the purpose of the study so may have made extra effort

towards accurate diagnosis. Additionally, studies varied on the cut-off scores used to define distress.

This evidence supports the recommendation in NICE CG123 that clinicians who are not competent to perform mental health assessments should refer people with possible common mental health disorders to appropriate healthcare professionals.

**Key reference**

Mitchell AJ, Rao S, Vaze A (2011) Can general practitioners identify people with distress and mild depression? A meta-analysis of clinical accuracy. Journal of Affective Disorders 130: 26–36

### Validation of the Geriatric Depression Scale

NICE CG123 recommends that when assessing a person with a suspected common mental health disorder, clinicians should consider using: a diagnostic or problem identification tool or algorithm, for example, the Improving Access to Psychological Therapies (IAPT) screening prompts tool; or a validated measure relevant to the disorder or problem being assessed, for example, PHQ-9, the Hospital Anxiety and Depression Scale (HADS) or GAD-7 to inform the assessment and support the evaluation of any intervention.

**Mitchell et al. (2010)** did a meta-analysis of the accuracy and clinical utility of the Geriatric Depression Scale (GDS) to detect depression in primary care, and to identify its value over routine identification of depression. Studies assessing the diagnostic validity of the GDS in people of 55 years or older were included. To calculate the value of the GDS, the unassisted diagnostic ability of GPs for detecting depression was also assessed.

Overall accuracy was calculated as the proportion of all cases that were either true positives or true negatives (known as the fraction correct or efficiency). A fraction correct of 60% or more was considered 'adequate' and 80% or more was considered 'good'. The clinical utility index was also calculated: positive clinical utility is the sensitivity multiplied by the positive predictive value; the negative clinical utility is the specificity multiplied by the negative predictive value.

In meta-analysis of 7 studies (n=1762) using $GDS_{30}$, the sensitivity was 77.4% (95% CI 66.3% to 86.8%) and specificity was 65.4% (44.2% to 83.8%), with a fraction correctly identified of 71.2% (95% CI 68.9% to 73.3%). The clinical utility was assessed as 'poor' for case finding (positive utility=0.29) and 'adequate' for screening (negative utility 0.60). In 10 studies (n=3012) of $GDS_{15}$ the sensitivity was 81.3% (95% CI 77.2% to 85.2%) and specificity was 78.4% (71.2% to 84.8%), with a fraction correctly identified of 77.6% (95% CI 77.5% to 80.4%). The clinical utility was assessed as 'poor' for case finding (positive utility=0.32) and 'good' for screening (negative utility 0.75).

$GDS_{15}$ was assessed as better than clinician detection alone, with 4% more correct positive diagnoses and 4% more correct negative diagnoses. $GDS_{30}$ was assessed as poorer than clinician detection alone. The authors additionally noted that the shorter version, $GDS_{15}$, may be more acceptable in primary care. The authors did not discuss limitations of their meta-analysis, but did state that the reasons that $GDS_{15}$ may be better than $GDS_{30}$ are uncertain.

Although the evidence suggests that $GDS_{15}$ may be useful for assessing possible depression in older people in primary care, this study does not show whether $GDS_{15}$ has any benefit over the instruments suggested in NICE CG123 in this population. Therefore, this evidence is not likely to affect current guidance.

Additional information about the study by Mitchell et al. (2010) is also available from an independent critical appraisal report produced for the Centre for Reviews and Dissemination's Database of Abstracts of Reviews of Effects.

**Key reference**

Mitchell AJ, Bird V, Rizzo M et al. (2010) Diagnostic validity and added value of the geriatric depression scale for depression in primary care: a meta-analysis of $GDS_{30}$ and $GDS_{15}$. Journal of Affective Disorders 125:10–17

**Supporting reference**

Centre for Reviews and Dissemination (2011) Diagnostic validity and added value of the geriatric depression scale for depression in primary care: a meta-analysis of $GDS_{30}$ and $GDS_{15}$. Database of Abstracts of Reviews of Effects

### PHQ-9 cut-off scores

NICE CG123 recommends that when assessing a person with a suspected common mental health disorder, consider using: a diagnostic or problem identification tool or algorithm, for example, the IAPT screening prompts tool; or a validated measure relevant to the disorder or problem being assessed, for example, PHQ-9, the HADS or GAD-7 to inform the assessment and support the evaluation of any intervention.

**Manea et al. (2012)** conducted a meta-analysis of 18 cross-sectional validation studies of PHQ-9 (n=7180) that defined major depressive disorder according to ICD or DSM criteria and provided sufficient data to allow calculation of contingency tables. Studies were excluded if diagnoses were not made with a standardised diagnostic interview (for example the Mini International Neuropsychiatric Interview [MINI]). 8 of the included studies validated PHQ-9 in primary care.

The pooled sensitivity for a cut-off score of 10 was 85% (95% CI 75% to 91%) and the pooled specificity was 89% (95%CI 83% to 92%). At a cut-off of 10, the sensitivity in primary care (89%, 95% CI 66% to 97%) was higher than that for hospital settings (74%, 95% CI 55% to 86%) but the specificity was similar (88% for primary care, 95% CI 80% to 93% vs 89% for hospital settings, 95% CI 66% to 97%). However, no data were reported for the significance of these differences between primary care and hospital settings.

A high level of heterogeneity was noted between studies, but a subsequent meta-regression analysis showed that blinding in applying the reference test was the only source of heterogeneity that affected the result (p=0.032). The diagnostic performance was not affected by the proportion of women (p=0.39), study setting (p=0.73), prevalence of depression (p=0.70), sample size (p=1.00) or mean age (p=0.28).

For other cut-off scores the specificity increased with the cut-off score, ranging from 73% (95% CI 63% to 82%) at a cut-off of 7 to 96% (95% CI 94% to 97%) at a cut-off of 15. Sensitivity did not decrease uniformly with increase in cut-off score.

The authors noted several potential limitations to their analysis including that publication bias could not be ruled out, that only one author selected studies for inclusion and that the quality assessment criteria used had not been validated. In 14 studies, participants were assessed by both PHQ-9 and the reference test; the other 4 used varying criteria from the results of PHQ-9 to determine whether the reference test was also used, which could have led to an overestimate of sensitivity.

This evidence suggests that PHQ-9 is a useful instrument for diagnosing depression, but that a single cut-off score might not be appropriate in all settings. In general practice, a cut-off score of 10 may be useful. This evidence is consistent with the use of PHQ-9 as an example of a validated instrument to use in assessment of common mental health disorders as recommended in NICE CG123.

Additional information about the study by Manea et al. (2012) is also available from an independent critical appraisal report produced for the Centre for Reviews and Dissemination's Database of Abstracts of Reviews of Effects.

**Key reference**

Manea L, Gilbody S, McMillan D (2012) Optimal cut-off score for diagnosing depression with the patient health questionnaire (PHQ-9): a meta-analysis. CMAJ 184: DOI:10.1503/cmaj.110829

**Supporting reference**

Centre for Reviews and Dissemination (2012) Optimal cut-off score for diagnosing depression with the patient health questionnaire (PHQ-9): a meta-analysis. Database of Abstracts of Reviews of Effects

## 1.4 Steps 2 and 3: treatment and referral for treatment

### Telephone counselling in people with acquired physical disabilities

NICE CG123 does not include recommendations for telephone counselling.

**Dorstyn et al. (2011)** did a meta-analysis of 8 studies (n=658) of telephone counselling to facilitate psychosocial recovery in adults with an acquired physical disability (for example amputation, severe burns or stroke). Included studies used at least 4 telephone sessions versus a control group, used standardised psychological measures, and provided sufficient data to allow calculation of effect sizes. Studies were excluded: if they were of chronic medical conditions (for example hypertension) or cognitive impairments (for example traumatic brain injury) but did not have a physical disability; or if the telephone component did not have a psychological focus or if the counselling component was not delivered by telephone.

The telephone counselling studies identified for the meta-analysis included participants with multiple sclerosis (5 studies), spinal cord injury (2 studies), a mix of spinal cord injury, nervous system disease and stroke (1 study). No studies of telephone counselling in people with severe burns or limb amputations were found. An average of 8 (SD=3.4) telephone counselling sessions were provided with varying length of individual sessions (range 30–90 minutes). Frequency also varied between studies (weekly, fortnightly, or weekly tapering to monthly over time). Frequency and duration were generally tailored to the person's needs.

Telephone counselling had a significant immediate post-treatment effect on coping strategies (effect size=0.57, 95% CI 0.22 to 0.92). A significant effect immediate post-treatment was also seen for community integration (effect size=0.45, 95% CI 0.19 to 0.71), and programmes to manage clinical depression (effect size=0.44, 95% CI 0.20 to 0.70).

The authors discussed several potential limitations of their meta-analysis including that telephone counselling is under-researched, particularly in people with physical disabilities and that most treatment effect analyses were done for single studies, which may be less reliable than using multiple studies. Additionally some studies had high attrition rates but did not use intention-to-treat analyses, which may have led to overestimation of effect sizes.

This study shows that the evidence for telephone counselling in people with acquired physical disabilities is limited, but may indicate beneficial effects on coping strategies, community integration and depression. Current evidence is not likely to affect NICE CG123.

Additional information about the study by Dorstyn et al. (2011) is also available from an independent critical appraisal report produced for the Centre for Reviews and Dissemination's Database of Abstracts of Reviews of Effects.

**Key reference**

Dorstyn DS, Mathias JL, Denson LA (2011) Psychosocial outcomes of telephone-based counseling for adults with an acquired physical disability: a meta-analysis. Rehabilitation Psychology 56: 1–14

**Supporting reference**

Centre for Reviews and Dissemination (2011) Psychosocial outcomes of telephone-based counseling for adults with an acquired physical disability: a meta-analysis. Database of Abstracts of Reviews of Effects

## 1.5 Developing local care pathways

No new key evidence was found for this section.

# 2 New evidence uncertainties

No new evidence uncertainties were identified during the Evidence Update process, however any uncertainties that may be identified in future for common mental health disorders will be added to the NHS Evidence UK Database of Uncertainties about the Effects of Treatments (DUETs). Other uncertainties can be found in the NICE research recommendations database.

UK DUETs was established to publish uncertainties about the effects of treatments that cannot currently be answered by referring to reliable up-to-date systematic reviews of existing research evidence.

# Appendix A: Methodology

## Scope

The scope of this Evidence Update is taken from the scope of the reference guidance:

- Common mental health disorders. NICE clinical guideline 123 (2011).

## Searches

The literature was searched to identify studies and reviews relevant to the scope. Searches were conducted of the following databases, covering the dates 01 September 2010 (the end of the search period of NICE clinical guideline 123) to 15 October 2012:

- CDSR (Cochrane Database of Systematic Reviews)
- CENTRAL (Cochrane Central Register of Controlled Trials)
- CINAHL (Cumulative Index to Nursing and Allied Health Literature)
- DARE (Database of Abstracts of Reviews of Effects)
- EMBASE (Excerpta Medica database)
- HTA (Health Technology Assessment) database
- MEDLINE (Medical Literature Analysis and Retrieval System Online)
- NHS EED (Economic Evaluation Database)
- PsycINFO

Table 1 provides details of the MEDLINE search strategy used, which was adapted to search the other databases listed above. The search strategy used in the reference guidance was modified to provide a more focused set of results; this was tested to ensure that the comprehensiveness of the results was not compromised. The search strategy was used in conjunction with validated Scottish Intercollegiate Guidelines Network search filters for RCTs and systematic reviews.

Figure 1 provides details of the evidence selection process. The long list of evidence excluded after review by the Chair of the EUAG, and the full search strategies, are available on request from contactus@evidence.nhs.uk

There is more information about how NICE Evidence Updates are developed on the NHS Evidence website.

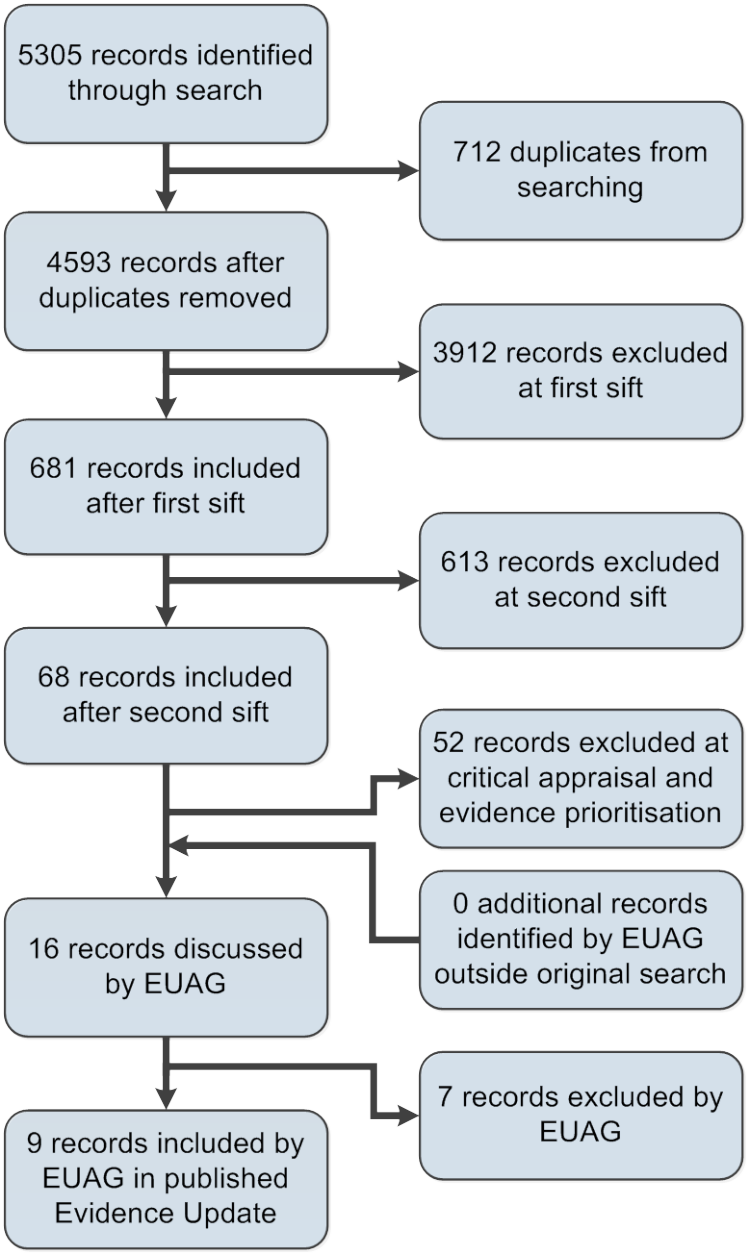
**Table 1 MEDLINE search strategy (adapted for individual databases)**

| | |
|---|---|
| **1** | exp *anxiety/ or exp *anxiety disorders/ or *agoraphobia/ or *neurocirculatory asthenia/ or exp *obsessive-compulsive disorder/ or *obsessive hoarding/ or *panic disorder/ or *phobic disorders/ or exp *stress disorders, traumatic/ or *combat disorders/ or *stress disorders, post-traumatic/ or *stress disorders, traumatic, acute/ or exp *depressive disorder/ or *depression/ |
| **2** | exp *Body Dysmorphic Disorders/ |
| **3** | exp *Compulsive Behavior/ |
| **4** | exp *Panic/ |
| **5** | exp *Stress, Psychological/ |
| **6** | (anx$ or panic$ or phobi$ or agoraphobi$ or gad or posttrauma$ or post trauma$ or ptsd or ocd or depress$ or dysthym$ or dysphor$ or dysmorph$ or obsessi$ or compulsi$ or seasonal affective disorder$ or (psych$ adj (stress or trauma$))).ti,ab. |
| **7** | 1 or 2 or 3 or 4 or 5 or 6 |
| **8** | health services accessibility/ or healthcare disparities/ |
| **9** | Health Status Disparities/ |
| **10** | exp health promotion/ and (access$ or barrier$ or disparit$ or equity or inequit$ or inequalit$).ti,ab. |
| **11** | ((access$ or barrier$ or disparit$ or equity or inequit$ or inequalit$ or uptake) adj4 ((health adj (care or service)) or healthcare)).ti,ab. |
| **12** | 8 or 9 or 10 or 11 |
| **13** | *checklist/ or *geriatric assessment/ or *interview/ or *interview, psychological/ or *needs assessment/ or *nursing assessment/ or *"outcome and process assessment (health care)"/ or *"outcome assessment (health care)"/ or exp *personality assessment/ or *"predictive value of tests"/ or exp *psychiatric status rating scales/ or exp *psychological tests/ or *questionnaires/ or *risk assessment/ |
| **14** | exp *diagnosis/ or *nursing diagnosis/ |
| **15** | (risk$ adj2 assess$).ti,ab. |
| **16** | (risk$ and assess$ and (diagnos$ or instrument$ or interview$ or inventor$ or item$1 or measur$ or psychometr$ or questionnaire$ or scale$ or score$ or screen$ or self report$ or subscale$ or sub$ scale$ or test form$ or tool$1)).ti,ab. |
| **17** | 13 or 14 or 15 or 16 |
| **18** | *"predictive value of tests"/ or exp *psychiatric status rating scales/ or exp *psychological tests/ or *questionnaires/ or exp *diagnosis/ or *nursing diagnosis/ |
| **19** | ((index or instrument$ or interview$ or inventor$ or item$ or measure$1 or questionnaire$ or scale$ or score$ or screen$ or self report$ or subscale$ or survey$ or tool$ or test form$) adj5 (detect$ or diagnos$ or identif$ or predict$ or psychodiagnos$ or recogni$)).ti,ab. |
| **20** | (area under curve or predictive value of tests or reproducibility of results or roc curve or validation studies or (sensitivity and specificity)).sh. |
| **21** | (accurac$ or accurat$ or area under curve or auc value$ or (likelihood adj3 ratio$) or (diagnostic adj2 odds ratio$) or ((pretest or pre test or posttest or post test) adj2 probabilit$) or (predict$ adj3 value$) or receiver operating characteristic or (roc adj2 curv$) or reliabil$ or sensititiv$ or specificit$ or valid$).ti,ab. |
| **22** | 18 or 19 |
| **23** | 20 or 21 |
| **24** | 22 and 23 |
| **25** | critical pathways.sh. |
| **26** | (models, nursing or models, organizational or models, theoretical).sh. and (care or healthcare or health care or service$).hw. |
| **27** | delivery of health care, integrated/ |
| **28** | exp Patient care management/ |
| **29** | ((care or interven$ or healthcare or intervention$ or program$ or therap$ or treat$ or organi?ation$ or system$) and (model$1 or path way$ or pathway$) and (collaborat$ or cooperative$ or co operative$ or coordinat$ or co ordinat$ or interdisciplin$ or inter disciplin$ or interprofessional or inter professional or liaison or multidisciplin$ or multi disciplin$ or multiprofessional$ or multi professional$ or team$1 or |

| | |
|---|---|
| | teamwork$)).ti,ab. |
| **30** | (stepped care or triage or case management).ti,ab. |
| **31** | (care adj3 organi$).ti,ab. |
| **32** | 25 or 26 or 27 or 28 or 29 or 30 or 31 |
| **33** | *"predictive value of tests"/ or exp *prognosis/ |
| **34** | (predict$ or prognos$).ti,ab. |
| **35** | *risk/ or *risk assessment/ or *Risk Factors/ |
| **36** | 33 or 34 or 35 |
| **37** | *medication adherence/ or *patient compliance/ or *patient dropouts/ or *patient readmission/ or *recurrence/ or |

| | |
|---|---|
| | *remission/ or *treatment failure/ or *treatment outcome/ or *treatment refusal/ |
| **38** | (adher$ or complian$ or dropout$ or readmi$ or recur$ or remi$ or (treatment adj2 failure) or (treatment adj2 outcome$) or (treatment adj2 refus$)).ti,ab. |
| **39** | 37 or 38 |
| **40** | 36 and 39 |
| **41** | 12 or 17 or 24 or 32 or 40 |
| **42** | 7 and 41 |

**Figure 1 Flow chart of the evidence selection process**




EUAG – Evidence Update Advisory Group

# Appendix B: The Evidence Update Advisory Group and Evidence Update project team

## Evidence Update Advisory Group

The Evidence Update Advisory Group is a group of topic experts who review the prioritised evidence obtained from the literature search and provide the commentary for the Evidence Update.

**Professor Tony Kendrick – Chair**

Professor of Primary Care, Dean, Hull York Medical School

**Mr Mike Bessant**

Regional Mental Health Lead, NHS Direct, Bristol

**Dr David Ekers**

Senior Clinical Lecturer in Psychological Interventions, School of Medicine Pharmacy and Health, Durham University

**Professor Linda Gask**

Professor of Primary Care Psychiatry, University of Manchester

**Professor Simon Gilbody**

Director of York Mental Health & Addictions Research Group, University of York, and Professor of Mental Health Services Research, Hull York Medical School

**Dr Rupert Suckling**

Assistant Director of Public Health, NHS Doncaster

## Evidence Update project team

**Marion Spring**

Associate Director

**Chris Weiner**

Consultant Clinical and Public Health Adviser

**Cath White**

Programme Manager

**Lynne Kincaid**

Medical Writer

**Bazian**

Information specialist support